UNIVERSITY OF
ILLINOIS LIBRARY
AT URBANA-CHAMPAIGN
STACKS

Digitized by the Internet Archive
in 2011 with funding from
University of Illinois Urbana-Champaign

http://www.archive.org/details/whatdozielskesre529simo

22 pages

Faculty Working Papers

WHAT DO ZIELSKE'S REAL DATA REALLY SHOW ABOUT PULSING?

Julian L. Simon , Professor, Departments of Business Administration and Economics

#529

College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

November 16, 1978

WHAT DO ZIELSKE'S REAL DATA REALLY SHOW ABOUT PULSING?

Julian L. Simon , Professor, Departments of Business Administration and Economics

#529

Summary:

A spaced advertising schedule is much more dollar-effective than is a massed (pulsed) schedule, according to this re-analysis of the experimental data developed by Pomerance and Zielske. This finding is at variance with the conclusions drawn by previous writers. The raw Pomerance-Zielske data are published here for the first time; until now only idealized representations of the data have been available, which has been the source of confusion and misunderstanding.

Acknowledgment:

At the University of Illinois Marketing Workshop in February, 1977 the participants threw valuable suggestions and criticisms at me so fast that I was hard put to make notes of the ideas let alone the names of those who offered the ideas. Alan Andreasen, Dave Gardner, Vithala Rao, Greg Upah, and Rick Winter were among those whome I remember clearly, but I am grateful to all. I also appreciate a searching critique by Edward C. Strong.

INTRODUCTION

The subject of the research reported in Zielske's well-known article (1959) is the comparison of massed advertising ("flighting" or "pulsing") versus spaced advertising schedules. This is made clear by the introductory statement in the less-well-known earlier report on the same work by Pomerance and Zielske (P-Z hereafter):

> Decisions have to be made continuously on how advertising should be scheduled. For instance, should it be concentrated in an intensive burst, or should it be spread over a longer period? Should a relatively small group of people be exposed many times, or a larger group a smaller number of times?
>
> In other words, what is the rate at which consumers can be made to remember advertising, and the rate at which they forget it? (Pomerance and Zielske, 1958)

The conclusion is hedged in both the P-Z and Zielske pieces. They say that a pulsed schedule will produce a higher peak number of rememberers whereas a spaced schedule will produce a higher average number of rememberers.

The ultimate business goal, however, is profit rather than peaks of remembrance. Or with respect to a given expenditure for advertising, the goal is to maximize the dollars of sales. Other measures of advertising are simply proxies for sales, and ultimately they must be related to sales. The issue, then, is whether massed or spaced advertising produces more sales per dollar of advertising.

If we consider recall of advertisements as a proxy for sales--always a questionable assumption, but an assumption that underlies the bulk of advertising research--then comparison of the average recalls during the entire relevant period indicates the relative selling power of two equal-cost campaigns. By Zielske's calculation, monthly spaced advertising produces 29 percent average recall over the year compared to

21 percent for the thirteen weekly advertisements (p. 241), apparently an advantage for spaced advertising. But Strong's simulation, based on the P-Z data as well as additional data of his own, yields estimates of 30.1 percent and 26.5 percent respectively, a somewhat smaller advantage for spaced advertising. And, Strong--using data he collected from a bi-weekly schedule, and the P-Z data plus some of his own monthly-exposure data--concluded from a simulation that a schedule of four bi-weekly flights of 4, 3, 3, and 3 exposures spaced throughout the year "delivers an average annual recall of 32%, two percentage points greater than that delivered by the monthly schedule" (1974, p. 377), i.e., an advantage for pulsing. Rao (1970), and Ackhoff and Emshoff (1975a; 1975b) have also asserted, and the proposition apparently is generally accepted, that in some circumstances pulsing can produce more sales dollars per dollar of advertising than spaced advertising.

My re-analysis in this paper of the P-Z data suggests that pulsing is never worthwhile--at least for the conditions of the P-Z experiment. Furthermore, this re-analysis shows a massive advantage for spaced advertising--selling power twice as great--which is a much greater advantage than suggested by either the Zielske or the Strong methods of calculation of average recall.

In addition to the finding that spaced advertising is much more dollar-effective than is massed advertising, this article presents for the first time the raw P-Z data, wherein lies an instructive tale of scientific history which is told in Appendix B. The very important and imaginative study that produced the data mentioned in the title was directed by Eugene Pomerance and Hugh (sic) Zielske (Pomerance and

include a re-analysis of the raw P-Z data. And I had just become aware of Stewart's massive study (1964) that used the P-Z design but that reproduced Zielske's diagrams as if they represented actual data points.

I therefore wrote to Zielske as follows (July 13, 1965):

> I would like to write a note for the Journal of Marketing commenting on the shortcomings of your presentation of the data. I could best do so if your actual data were shown. But of course the data are yours and I assume were sent to me in confidence.
>
> So two questions: 1) Would you like to co-author the note with me and include the data? or, 2) May I hav

Zielske replied that he "would have no interest in co-authoring a 'note'" and he refused me permission to quote his data, on the grounds that I misquoted him in my 1965 article. He added

> I am fully aware that all people who have quoted [my] article have not also always quoted the qualifications or the information that was included on methodology. But whether or not more technical detail would have reduced this is very doubtful.
>
> A full technical description of how the data were processed, together with the raw data, has been available and distributed to all who requested such information.
>
> At the time the article was submitted for publication, back in 1958, the Journal felt (correctly, I believe) that there would be insufficient interest in this "technical appendix" to warrant the space it would require, and it's probably safe to assume that interest would be even less now--seven years later."

Whether rightly or wrongly I do not know, but I accepted Zielske's refusal and did nothing in print about this matter. Recently, however, I (with Arndt) began to prepare a new article on the advertising response function and economies of scale, and again my thoughts turned to the P-Z

data. I asked--and this time received--permission from Eugene Pomerance, for Foote, Cone and Belding, Inc., to reproduce the raw data. Afterwards our literature search revealed that Strong (1974; see also Strong, 1972) replicated the P-Z study, and included a graphic presentation of the original P-Z data in his report.* As noted earlier, Strong's data confirm the P-Z data where they overlap; this is comforting, and strengthens the substantive analysis presented in the article.

That the P-Z data have not been in print is more than regrettable. Considerable misunderstanding about the substantive issue of spaced versus massed advertising has been caused. Furthermore, as Strong put it "The accretion of wisdom on top of Zielske's admittedly brilliant work has been severely hampered by the ill-advised smoothing he performed" (letter of March 8, 1978).

---

*Another element in what would be a comedy if it were funny is that when I wrote Dr. Edward Strong (sic) at Stanford asking to see a copy of his thesis, I received a reply that Dr. Strong had died in 1963. This made me wonder whether my belief that there was a contemporary Edward Strong was in error, and the confusion was compounded in a MSI working paper by Ray (1975) that lists the 1914 publication of Edward K. Strong under the same author heading as the 1972 thesis of Edward C. Strong.

Zielske, 1958), though the attribution is invariably to Zielske as published in Zielske (1959). In neither the original articles are the raw data shown, but rather smoothed idealized graphs of them are given. Since then many studies have referred to the Zielske article, and have treated the smoothed figures as if they are actual data points. It is only the exception (e.g., Lodish, 1971; Corkindale and Kennedy, 1975, p. 80) which has mentioned that the figures are idealized and do not represent actual observations. Lodish also noted that he was unable to obtain the raw P-Z data for re-analysis, and therefore used the idealized representations for his empirical work. In 1963 I obtained the raw data from an internal publication of Foote, Cone, and Belding, but I was not able to obtain permission to publish them until now.

THE METHOD AND DATA

The main details of the P-Z method (for more general background see Zielske, 1959) are as stated in the internal research report of Foote, Cone and Belding (1958):

> A total of 5668 names was selected from the 1955 Chicago telephone directory. This total was subdivided into 26 test samples of 143 respondents each and 26 control samples of 75 respondents each....
>
> The following steps show in detail how the individual samples were selected:
>
> 1. 5668 names were systematically selected from the 1955 Chicago telephone directory by choosing approximately every 160th residential listing. The names, together with phone numbers and addresses, were typed on the questionnaires. All male names were preceded with "Mrs."
>
> 2. The questionnaires, which were kept in alphabetical sequence throughout the selection process, were then divided into 26 groups of 218 each. The first

questionnaire was placed in group 1, the second in group 2, etc., until all 26 groups were complete. Groups 1 through 12 were designated as the weekly exposure groups--test respondents in these groups received ads at weekly intervals. Groups 13 through 26 were designated as the four-week-interval exposure groups--test respondents in these groups received ads at intervals of four weeks.

3. From each of the 26 groups of 218 questionnaires, approximately every third one was selected as control respondents who received no ads but were interviewed at the same time as the test respondents. The purpose of the control samples was to see what the level of recall was among those not exposed to the ads. (There was no significant trend in remembrance of normally scheduled advertising for the product among the control groups.) There were 75 respondents in each control sample and 143 in each test sample.

These samples yielded an average of 91 completed interviews for each of the 26 test exposure groups and 51 completed interviews for each control group. The overall completion rate was 65%....

Thirteen different ads were used in this study. In order to minimize any effect ad sequence might have on the results, the order in which the ads were mailed to each group of respondents was varied.

To help insure respondents' opening the letters, six different colored envelopes with personal return addresses were used.

All interviews were made by telephone, and a maximum of five calls were made in an attempt to contact each respondent.

The exposure schedules, interview schedules, and results for the weekly and every-four-weeks exposure groups were as shown in Table 1. The results are the recall percentages shown in the tables. These data are the points plotted in Figures 1 and 2 and Table 1.

Figures 1 and 2; Table 1

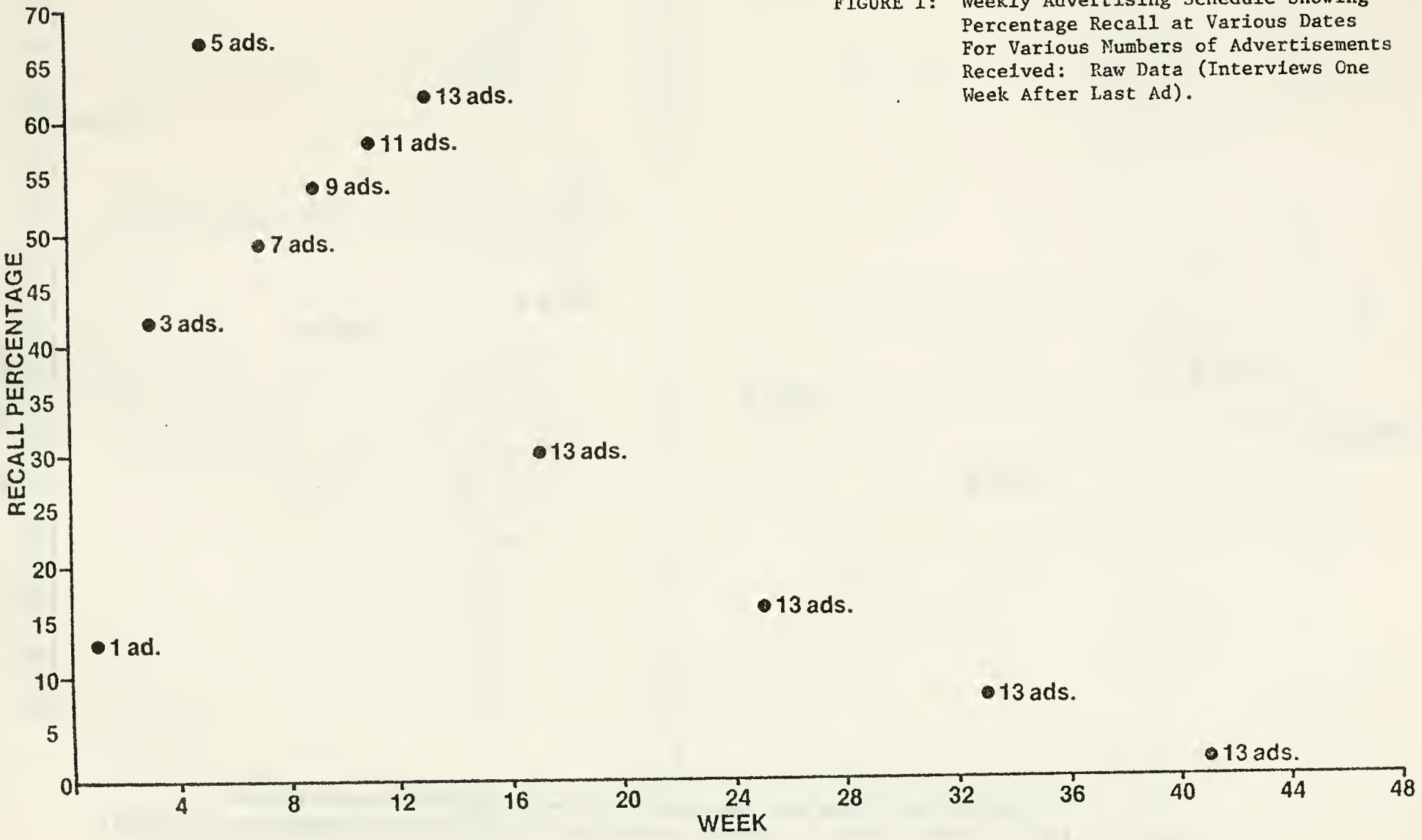


FIGURE 1: Weekly Advertising Schedule Showing Percentage Recall at Various Dates For Various Numbers of Advertisements Received: Raw Data (Interviews One Week After Last Ad).

FIGURE 2: Four-weekly Advertising Schedule Showing Recall at Various Dates for Various Numbers Of Advertisements Received: Raw Data (Interviews One Week After Last Ad).

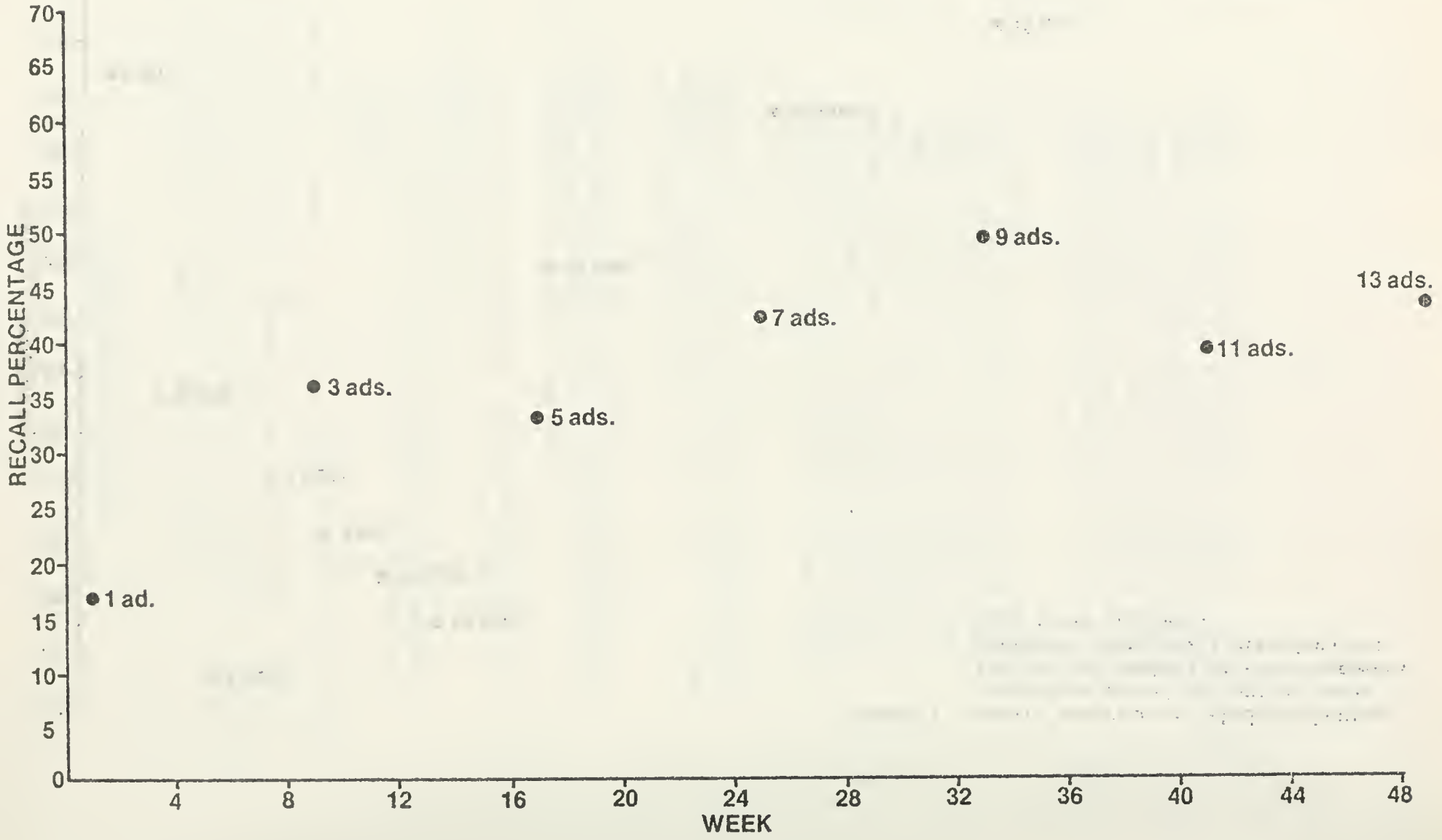


TABLE 1

Exposure schedule and the results for the groups that received advertisements weekly.

Group	Number of Ads Received At Weekly Intervals	Number of Weeks Between Mailing Of First Ad to Group and Start of Interviewing	Number of Weeks Between Mailing of Last Ad to Group and Start of Interviewing	Recall Percentage
1	1	1	1	13
2	3	3	1	42
3	5	5	1	67*
4	7	7	1	49
5	9	9	1	54
6	11	11	1	58
7	13	13	1	62
8	13	17	5	30
9	13	25	13	16
10	13	33	21	8
11	13	41	29	2
12	13	49	37	?

Exposure schedule and results for the groups that received advertisements every four weeks.

Group	Number of Ads Received At Four-Week Intervals	Number of Weeks Between Mailing of First Ad to Group and Start of Interviewing	Number of Weeks Between Mailing of Last Ad to Group and Start of Interviewing	Percentage Decrease in Recall	Recall Percentage
13	1	1	1		17
14	1	5	5	100	0
15	3	9	1		36
16	3	13	5	72	10
17	5	17	1		33
18	5	21	5	21*	26
19	7	25	1		42
20	7	29	5	49	21
21	9	33	1		49
22	9	37	5	40	29
23	11	41	1		39
24	11	45	5	33	26
25	13	49	1		43
26	13	53	5	27	31

*Measurement that was discarded by Zielske .

I treated the data in the following fashion:

1. The raw recall percentages for observations a week after the last advertising exposure, for the weekly and the every-four-weeks exposure schedules, were plotted as in Figures 1 and 2, taken from Table 1. For the four-week-exposure group, I also plotted the raw recall percentages for observations five weeks after the last advertising exposure (computed from the raw "decrease in recall" data in Figure 3.

Figure 3

2. The data for decrease in recall from one week after the last exposure to five weeks after the last exposure, for the every-four-weeks exposure groups, are plotted in Figure 4, and are smoothed by eye. This curve could be fitted analytically, but the choice of function could not conceivably affect the conclusions drawn below. Nor is there compelling scientific logic in favor of analytical functions in preference to eye-and-hand functions. Furthermore, it would not be feasible to fit the other curves in the paper analytically. But again, the conclusions are not sensitive to the exactness of the drawing, so there is no real problem.

Figure 4

3a. Smoothed curves for the proportions of the weekly exposure groups recalling the campaign were drawn by eye in Figure 5. The week-to-week forgetting during the 13-week exposure period was ignored because of its small magnitude.

Figure 5

3b. Smoothed curves were drawn for the proportions of the every-four-weeks groups remembering the campaign. The procedure was as follows:

Figure 3: Decrease in Recall During the Period From One to Five Weeks Following Various Series of Advertisements: Raw Data

Figure 4: Decrease in Recall During Four Weeks: Smoothed Curve

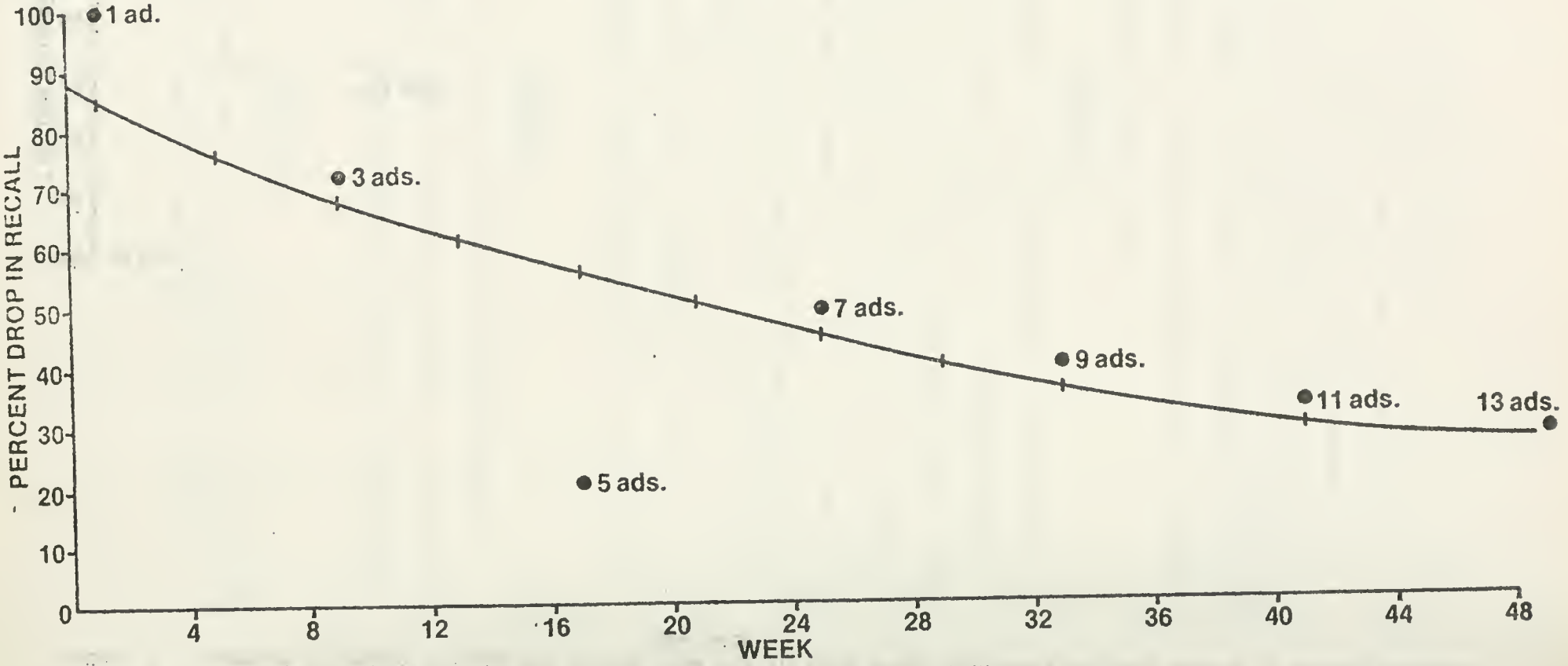


Figure 5: Weekly Advertisement Schedule: Smooth Curve Drawn by Eye, Bounding Approximately 1076 Recall Weeks.

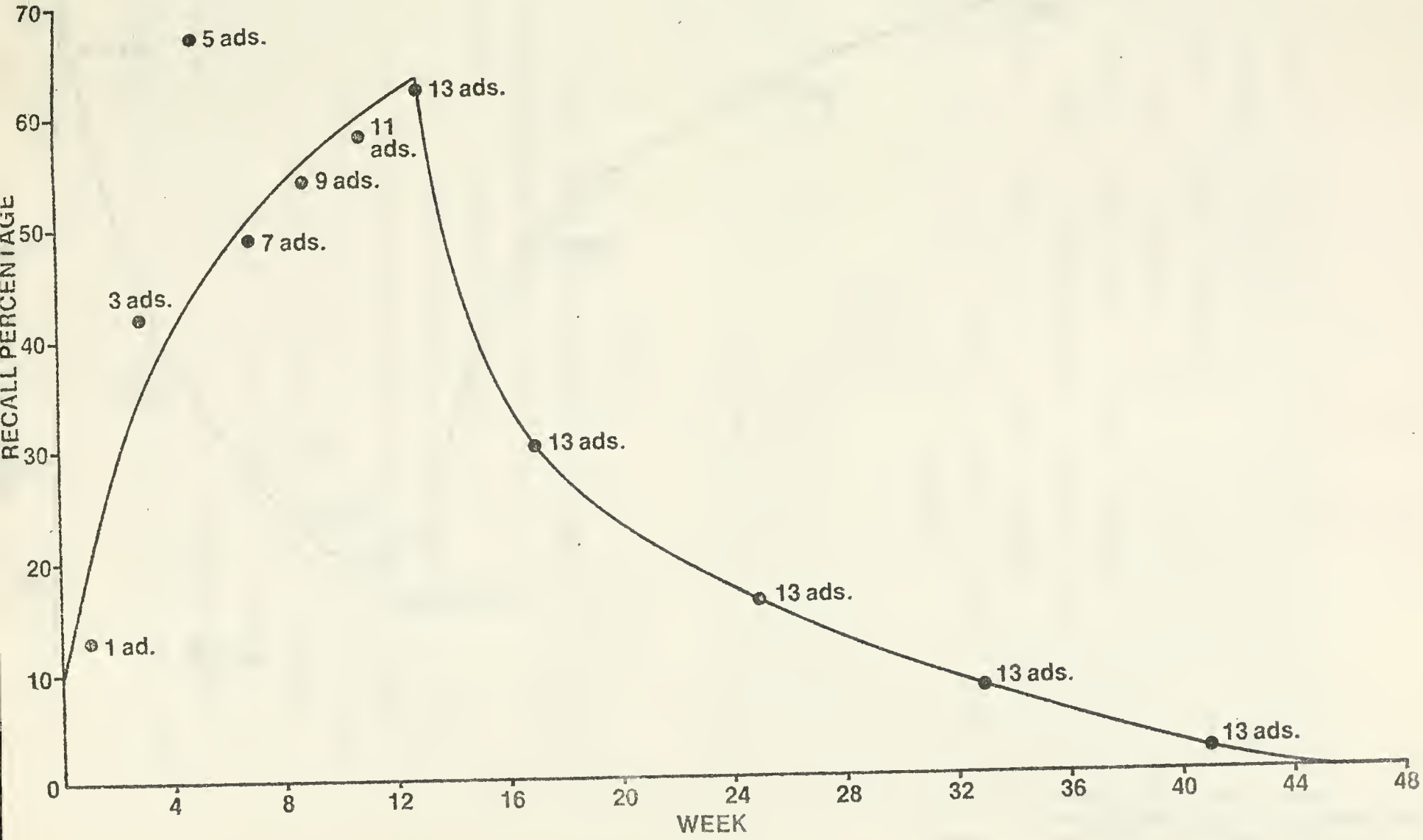
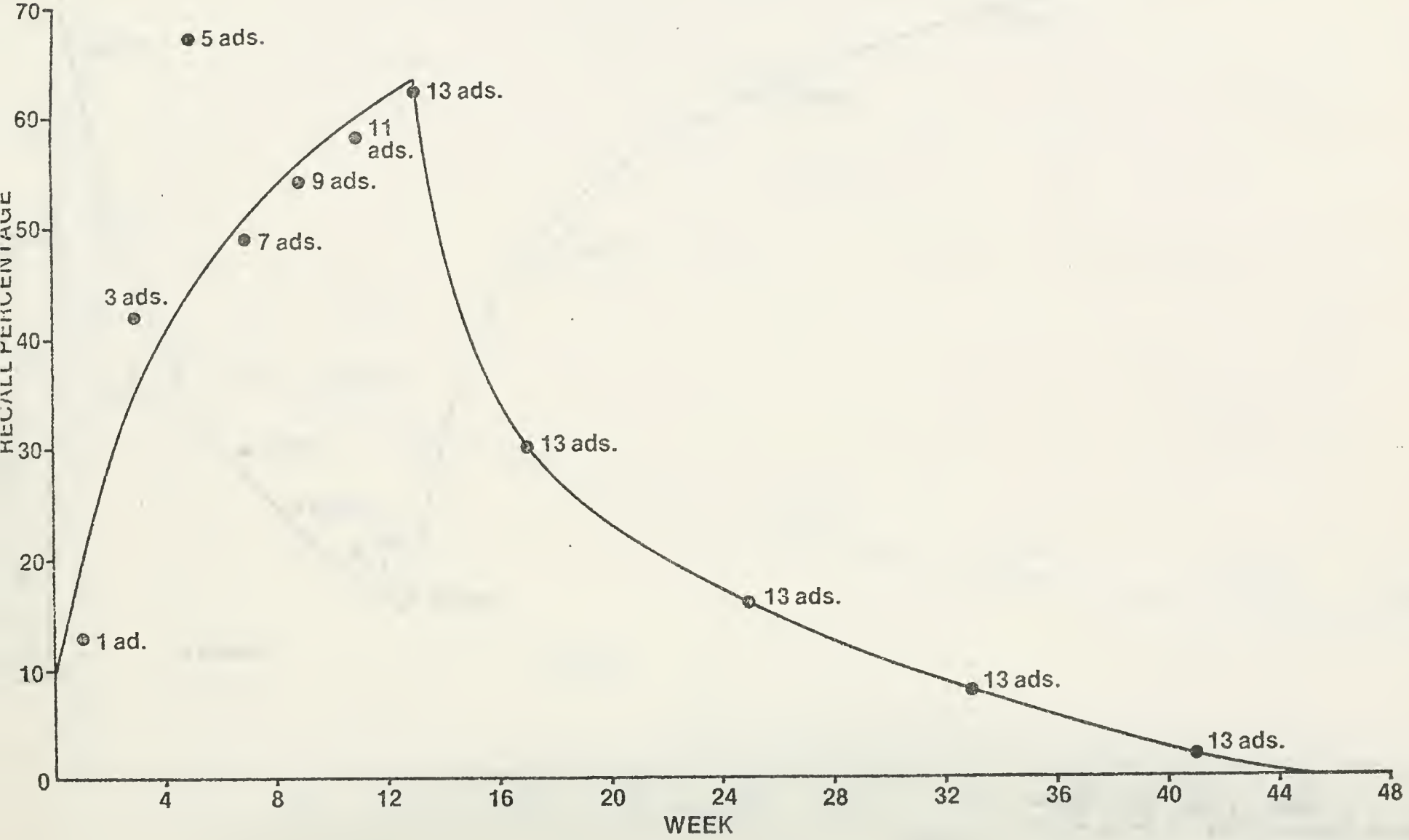


Figure 5: Weekly Advertisement Schedule: Smo[illegible]
Curve Drawn by Eye, Bounding Approxi-
mately 1076 Recall Weeks.



(i) Draw a smooth curve (the rising portion of the solid line in Figure 6) through the observations in Figure 2. (ii) Draw in the estimated recall points four weeks after each observation, e.g., 3% [(1-.85)(.20)] at week 5, and 10% [(1-.66)(.30)] at week 13, using the four-week decrease-in-recall function in Figure 4. (iii) Draw in the recall percentages estimated from the smoothed curves for unobserved numbers of exposures, using the smoothed envelope curve, e.g., 26% at 5 weeks, and 34% at 13 weeks. (iv) Draw in the recall percentages for those number of exposures four weeks later, e.g., 6% at 9 weeks. (v) Connect all the points in Figure 6.

Figure 6

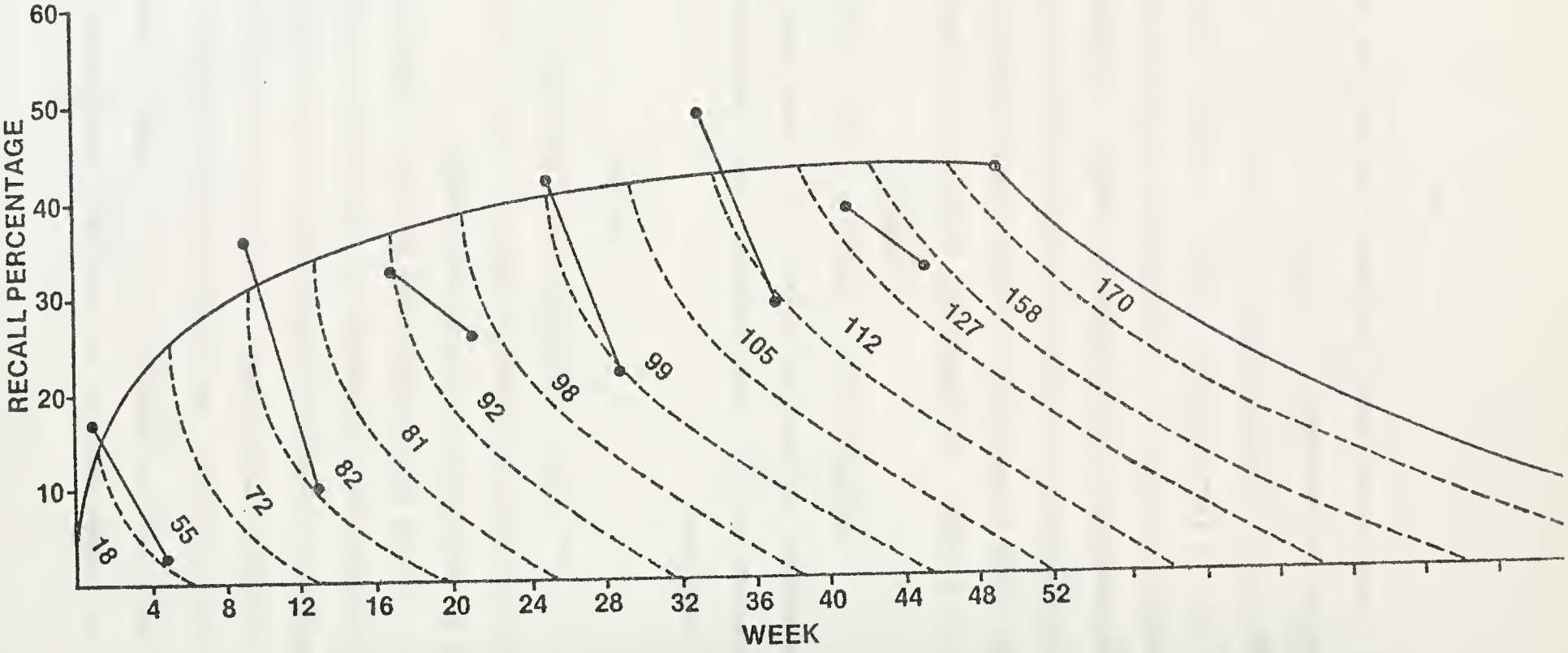
4. Calculate the area under the recall curves in Figures 5 and 6 by integration (counting the squares on the graph paper is a simple and effective device). This yields the total remembrances for the two types of campaigns.

RESULTS

1. The weekly schedule reaches a higher peak recall percentage than does the every-four-weeks schedule, as stated by P-Z (1958). But this is no surprise, nor of great importance.

2. The rate of forgetting the entire campaign after its completion is slower after 13 exposures spread over a year than after 13 exposures spread over 13 weeks as noted by P-Z; the decrease in recall from one to five weeks after the last of the thirteen ads is 70% for the weekly schedule, but only 29% for the four-weekly schedules. This means that though recall is higher for the weekly schedule immediately after the 13th exposure than for the every-four-weeks groups, after about 17 weeks

Figure 6: Four-weekly Advertising Schedule: Smoothed Curve of Recall



the rate of recall becomes higher and stays higher for the four-week group thereafter. (Compare Figure 6 to Figure 5.) But this is of more interest to the psychologist than to the marketing decision-maker.

3. The relevant measure of an advertising campaign's impact is the number of "recall weeks," that is, the number of weeks multiplied by the appropriate recall rate. This measure reveals the most important outcome of this study: the total amount of advertising impact (as measured by recall-weeks) is much higher for the every-four-weeks schedule than for the weekly schedule. For the every-four-weeks schedule there are roughly 2130 recall-weeks, compared with between 1054 and 1076 recall-weeks for the weekly campaign.

The implication is clear: A given advertising budget is most efficient if it is spread out over the maximum period rather than concentrated in a single burst. Several bursts will be better than one burst but worse than an even spread, by implication. This stark conclusion was not drawn by Pomerance and Zielske though it is implied in their finding about the higher average recall with the monthly schedule; and this conclusion requires the data and analysis shown here to substantiate it. It contradicts the assertion by Rao (1970, p. 55) but Rao presented no relevant data. It does not agree with Strong's result for a simulated comparison of bi-weekly flights spaced throughout the year versus monthly exposures, but this result of Strong's is several logical and assumptional steps removed from actually-observed data. And it should be remembered that where they overlap, Strong's raw data coincide nicely with the P-Z raw data.

DISCUSSION

The conclusions that may be drawn from Zielske's (and Strong's) data are only as general as the data themselves. The range of the data is limited as follows: (1) the advertised product was a food, "a staple in almost every home" (p. 1 of Foote, Cone and Belding research report). The results might be different for a durable or another sort of product that differs in interpurchase interval and in other ways. (2) The data were generated by just one advertising medium, direct mail, and direct mail would not be a likely medium for a staple food product. (3) The measure of advertising impact was unaided recall. A sales measure might show a different sort of response pattern. And a recall measure has a built-in saturation limit (100%) that is not relevant for actual purchases. Furthermore, the relationship of recall to sales might be different at different points in the marketing campaign. (4) All persons received advertisements starting in the same week. There might have been uncontrolled seasonal or other time effects that influenced the results.

These qualifications and criticisms are severe. But the response to them is strong: These are the only data there are. And until a wide range of data are produced, the results from these data must have some claim on our belief, especially as they have been so relied on for other purposes in the past. The fact that Strong's data, from an experiment run at another time and with a different product, fit well with Zielske's data, reduces concern about qualifications 1 and 4, however.

Zielske tested out only two patterns of advertising spacing, once-a-week for thirteen weeks and then no more advertising, and once-a-month for 13 months. (Strong also tested a bi-weekly schedule.) But there

are many other possible schedules, especially for new products, including a heavy dose of advertising to start with and then a lower advertising level. (This schedule is, in fact, shown deductively to be optimal by Levy and Simon, forthcoming.) Some pattern other than even spacing of advertisements might be better (once the equilibrium level for a new product is attained). Further experiments on a few of the most reasonable-seeming patterns would be necessary to make this possibility seem less likely.

The reader may wonder how these findings may be reconciled with the conclusion put forth by Rao (1970) and Ackhoff-Emshoff (1975a, p. 10) that pulsing can be better than spaced advertising. I cannot answer this question, because the raw data are shown in neither discussion, and no answer was received to my requests for the data.

SUMMARY AND CONCLUSIONS

The P-Z experiment pioneered a very useful method for understanding the effects of advertising. It is good that the raw data are finally available so that readers can know exactly what was done, and what the results were.

The most important finding is that for a given advertising budget and a given campaign time period, a campaign that spreads the exposures as widely as possible over the period produces much more sales impact (as measured in recall weeks) than does a pulsed campaign. And the advantage of the spaced campaign is very large in percentage terms. This present publication is the first solid direct evidence for this proposition.

APPENDIX: HOW THE P-Z DATA NOW APPEAR IN PRINT

The saga of how this article and the P-Z data finally appear in print is rather unusual, and I believe that it is illuminating and instructive. I hope the reader will understand the intrusion of such personal material into a professional article.

The first draft of this piece was written more than a decade ago, and the scotch tape on the manuscript has already cracked. In 1963, when teaching a course on research methods in advertising and communications--my first semester as a college teacher--I wished to use Zielske (1959) as a case study for class analysis. But when reading it closely so as to prepare an answer to one of my stock class questions--"What exactly did the author do?"--I found that I could not answer the question myself. So I wrote to Zielske, and in October, 1963 he sent me the raw data that underlie Table 1.

(Each time I have taught a research-methods course since then I have assigned the piece by Zielske near the beginning of the course. When I ask "What exactly did the author do?" the students at first give confident answers, but then there quickly appear contradictions between them, such as whether the total sample size was 3,650 or 13,000 or 15,000 or a variety of other suggestions. With as much sternness and drama as I can muster, I say "You mean you don't know how many groups there were? Better read it again for next week." And the next week, of course, the students are still confused, though the best ones say "The answers cannot be found in the article.")

A couple of years later I was revising a manuscript on the advertising response function and economies of scale, and I wished to

REFERENCES

Russell Ackoff and James R. Emshoff, "Advertising Research at Anheuser Busch, Inc. (1963-1968)," Sloan Management Review, Vol. 16, Winter 1975, pp. 1-15.

Russell Ackoff and James R. Emshoff, "Advertising Research at Anheuser Busch, Inc. (1968-1974)," Sloan Management Review, Vol. 16, Spring 1975, pp. 1-15.

Foote, Cone & Belding, "A Study of Advertising Scheduling, Internal Report," Chicago, 1958.

David R. Corkindale and Sherril H. Kennedy, Measuring the Effects of Advertising, (Lexington, Mass.: Saxon House/Lexington Books, 1975).

Haim Levy and Julian L. Simon, "Choosing the Best Advertising Appropriation When Appropriations Interact Over Time," in Jagdish Sheth (ed.), Research in Marketing, Vol. 1, forthcoming.

Leonard M. Lodish, "Empirical Studies on Individual Response to Exposure Patterns," Journal of Marketing Research, Vol. 8, May 1971, pp. 212-218.

Eugene Pomerance and Hubert (Hugh) Zielske, "How Frequently ShouLd You Advertise?" Media/Scope, September 1958.

Ambar G. Rao, Quantitative Theories in Advertising (New York: Wiley, 1970).

Julian L. Simon, "Are there Economies of Scale in Advertising?" Journal of Advertising Research, Vol. 5, June 1965, pp. 15-19.

Julian L. Simon, "Further Evidence on the Effect of Scale and Advertising," Journal of Advertising Research, Vol. 9, March 1969, pp. 38-41.

Julian L. Simon, Issues in the Economics of Advertising (Urbana: University of Illinois Press, 1970).

Edward C. Strong, "The Effects of Repetition in Advertising: A Field Experiment," Unpublished Ph.D. Dissertation, Stanford, 1972.

Edward C. Strong, "The Use of Field Experimental Observations in Estimating Advertising Recall," Journal of Marketing Research, Vol. 11, November 1974, pp. 369-378.

Edward K. Strong, "The Effect of Size of Advertisements and Frequency of their Presentation," Psychological Review, Vol. 21, March 1914, pp. 136-152.

Hubert (Hugh) Zielske, "The Remembering and Forgetting of Advertising," The Journal of Marketing, Vol. 23, January 1959, pp. 239-243.

BOUND BY
HM
3-94

UNIVERSITY OF ILLINOIS-URBANA
3 0112 060296404